REMARKS

The specification has been amended to correct typographical errors. In particular, the recitation of "C160S" and "P148A" have been corrected to "A160S" and "F148A," respectively. The present disclosure states that "[i]n the mutein nomenclature used herein, the changed amino acid is depicted with the native amino acid's one letter code first, followed by its position in the EPO molecule, followed by the replacement amino acid one letter code" (see specification as filed at p.15, ll.12-15). One of ordinary skill in the art would have knowledge of the native sequence of erythropoietin, which was published in, *e.g.*, Jacobs K. *et al.* 1985. "Isolation and characterization of genomic and cDNA clones of human erythropoietin," *Nature* 313(1985): 806-810, enclosed herewith as Exhibit A, and U.S. Patent Publication No. 2004-0122216, incorporated herein by reference (see, *e.g.*, p. 37, ¶¶ 0324-0326). Thus, the skilled artisan could deduce without difficulty that the native amino acids at positions 148 and 160 of EPO are F (phenylalanine) and A (alanine), respectively – not P (proline) and C (cysteine). Therefore, one of ordinary skill in the art would know that the muteins referred to in the present specification are F148A and A160S. Further, the recitation "prostrate" has been corrected to "prostate."

Claim 46 has been canceled without prejudice. Applicants reserve the right to prosecute the subject matter of the canceled claims in one or more related continuation, continuation-in-part or divisional applications. Claim 45 has been amended to delete the duplicate term K45D/R150E.

No new matter has been added. Upon entry of the present amendment claims 1, 37-45 and 47-69 will be pending.

CONCLUSION

Applicants respectfully request that the above-made amendments and remarks be entered and made of record in the present application.

No fee is believed to be required in connection with this amendment. However, should any fee be due, please charge the required amount to Jones Day Deposit Account No. 50-3013.

Respectfully submitted,

Date: January 10, 2008

Laura A. Coruzzi 30,742
Laura A. Coruzzi (Reg. No.)

JONES DAY
222 East 41st Street
New York, New York 10017
(212) 326-8383

by Aileen F. Falvey 46,097
Reg. No.

Received 12 October; accepted 12 December 1984.

1. Carswell, E. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 3666–3670 (1975).
2. Matthews, N. & Watkins, J. F. *Br. J. Cancer* **38**, 302–309 (1978).
3. Zacharchuk, C. M., Drysdale, B-E., Mayer, M. M. & Shin, H. S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6341–6345 (1983).
4. Hayashi, H., Kiyota, T., Sohmura, Y. & Haranaka, K. *Proc. 43rd Japan. Cancer Assoc. No. 1132*, 314 (1984).
5. Hori, K., Hayashi, H., Sohmura, Y. & Haranaka, K. *Proc. 43rd Japan. Cancer Assoc. No. 1130*, 314 (1984).
6. Matthews, N. *Immunology* **44**, 135–142 (1981).
7. Matthews, N. *Immunology* **48**, 321–327 (1983).
8. Williamson, B. D., Carswell, E. A., Rubin, B. Y., Prendergast, J. S. & Old, L. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5397–5401 (1983).
9. Blattner, F. R. *et al. Science* **196**, 161–169 (1977).
10. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157–1174 (1978).
11. Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).
12. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).
13. Thomas, M. & Davis, R. W. *J. molec. Biol.* **91**, 315–328 (1975).
14. Davis, R. W., Botstein, D. & Roth, J. R. (eds) *Advanced Bacterial Genetics*, 106–107 (Cold Spring Harbor, New York, 1980).
15. Messing, J. *Meth. Enzym.* **101**, 20–78 (1983).
16. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
17. Winzler, R. J. in *Hormonal Proteins and Peptides* (ed. Li, C. H.) 1–5 (Academic, New York, 1973).
18. Amann, E., Brosius, J. & Ptashne, M. *Gene* **25**, 167–178 (1983).
19. Corbett, T. H., Griswold, D. P., Roberts, B. J., Peckham, J. C. & Schabel, F. M. *Cancer Chemother. Rep. (Pt. 2)* **5**, 169–186 (1975).
20. Corbett, T. H., Griswold, D. P., Roberts, B. J., Peckham, J. C. & Schabel, F. M. *Cancer* **40**, 2660–2680 (1977).
21. Clark, I. A., Virelizier, J-L., Carswell, E. A. & Wood, P. R. *Infect. Immun.* **32**, 1058–1066 (1981).
22. Taverne, J., Dockrell, H. M. & Playfair, J. H. L. *Infect. Immun.* **33**, 83–89 (1981).
23. Taverne, J., Depledge, P. & Playfair, J. H. L. *Infect. Immun.* **37**, 927–934 (1982).
24. Playfair, J. H. L., Taverne, J. & Matthews, N. *Immun. Today* **5**, 165–166 (1984).
25. Vieira, J. & Messing, J. *Gene* **19**, 259–268 (1982).
26. Ito, H., Ike, Y., Ikuta, S. & Itakura, K. *Nucleic Acids Res.* **10**, 1755–1769 (1982).
27. Ruff, M. R. & Gifford, G. E. *J. Immun.* **125**, 1671–1677 (1980).
28. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
29. *Isoelectric Focussing: Principles and Methods (User's manual)* (Pharmacia Fine Chemicals, Sweden, 1982).

Isolation and characterization of genomic and cDNA clones of human erythropoietin

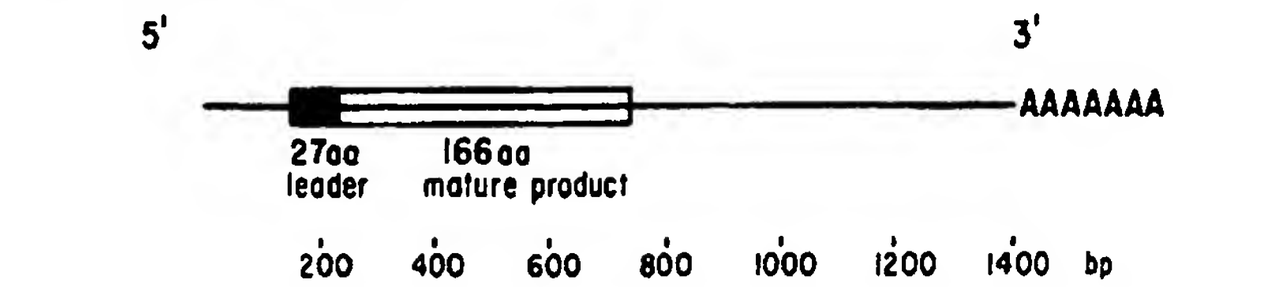
Kenneth Jacobs, Charles Shoemaker, Richard Rudersdorf, Suzanne D. Neill, Randal J. Kaufman, Allan Mufson, Jasbir Seehra, Simon S. Jones, Rodney Hewick, Edward F. Fritsch, Makoto Kawakita*, Tomoe Shimizu† & Takaji Miyake†

Genetics Institute, Inc., 225 Longwood Avenue, Boston, Massachusetts 02115, USA
* Kumamoto University, 39-1 Kurokami 2-Chome, Kumamoto-shi, 860 Japan
† Wright State University, Dayton, Ohio 45439, USA

The glycoprotein hormone erythropoietin regulates the level of oxygen in the blood by modulating the number of circulating erythrocytes, and is produced in the kidney[1–4] or liver[5,6] of adult and the liver[7,8] of fetal or neonatal mammals. Neither the precise cell types that produce erythropoietin nor the mechanisms by which the same or different cells measure the circulating oxygen concentration and consequently regulate erythropoietin production (for review see ref. 9) are known. Cells responsive to erythropoietin have been identified in the adult bone marrow[10], fetal liver[11] or adult spleen[12]. In cultures of erythropoietic progenitors, erythropoietin stimulates proliferation and differentiation to more mature red blood cells. Detailed molecular studies have been hampered, however, by the impurity and heterogeneity of target cell populations and the difficulty of obtaining significant quantities of the purified hormone. Highly purified erythropoietin may be useful in the treatment of various forms of anaemia, particularly in chronic renal failure[13–15]. Here we describe the cloning of the human erythropoietin gene and the expression of an erythropoietin cDNA clone in a transient mammalian expression system to yield a secreted product with biological activity.

Fig. 1 Northern analysis of human fetal liver mRNA. Human fetal liver (5 μg) and adult liver mRNA (5 μg) were electrophoresed in a 0.8% agarose/formaldehyde gel and transferred to nitrocellulose as described previously[41]. An erythropoietin-specific single-stranded probe was prepared from an M13 template containing the 87-bp exon of the human erythropoietin gene; the primer was a 20 mer derived from the same tryptic fragment as the original 17 mer probe. The ^{32}P-labelled probe was prepared as described previously[42] except that after digestion with *Sma*I, the small fragment was purified from the M13 template by chromatography on a Sepharose CL4b column in 0.1 M NaOH/0.2 M NaCl. The filter was hybridized to ~5×10^6 c.p.m. of this probe for 12 h at 68 °C, washed in 2×SSC at 68 °C and exposed for 6 days with an intensifying screen. Marker mRNAs of ~2,200 and 1,000 nucleotides (indicated by arrows) were run in an adjacent lane.

Methods. Erythropoietin was purified as described previously[27] except that the phenol treatment was eliminated and replaced by heat treatment at 80 °C for 5 min to inactivate neuraminidase and the final step in the purification was fractionation on a C-4 Vydac reverse-phase HPLC column (Separations Group) using a 0–95% acetonitrile gradient in 0.1% trifluroacetic acid (TFA) over 100 min. The position of erythropoietin in the gradient was determined by gel electrophoresis and N-terminal amino-acid sequence analysis[16] of the major peaks and comparing sequences obtained with those previously reported for erythropoietin[21–23]. Using this approach, erythropoietin was shown to elute at ~53% acetonitrile and represented 40% of the total eluted protein. Fractions containing erythropoietin were evaporated to ~100 μl, adjusted to *p*H 7 with 1 M ammonium bicarbonate and digested to completion with TPCK-treated trypsin (Worthington) (2% w/w enzyme/substrate) for 18 h at 37 °C. The tryptic digest was then subjected to reverse-phase HPLC using the conditions described above and the absorbance at both 280 and 214 nm monitored. Well-separated peaks were evaporated to near dryness and subjected directly to N-terminal sequence analysis[16] using an Applied Biosystems Model 470A gas phase sequenator. The sequences obtained are underlined in Fig. 2. Two of these tryptic fragments were chosen for synthesis of oligonucleotide probes. From the sequence Val-Asn-Phe-Tyr-Ala-Trp-Lys a 17 mer of 32-fold degeneracy (5'd(TTCCANGCG^TAG^AAG^TT); pool I) and a partially overlapping 18 mer of 128-fold degeneracy (5'd(CCANG CG^TAG^AAG^TTNAC); pool II) were prepared on an Applied Biosystems Model 380A DNA synthesizer. From the sequence Val-Tyr-Ser-Asn-Phe-Leu-Arg, two pools of 14 mers, each 48-fold degenerate (5'd(TAC$^{T}_{C}$A$^{T}_{C}$G$^{C}_{T}$T^{N}AA$^{T}_{C}$TTT^{C}CT); pool III) and 5'd(TAC$^{T}_{C}$A$^{T}_{C}$G$^{C}_{T}$T^{N}AA$^{T}_{C}$TTT^{C}TT); pool IV), which differ at the first position of the leucine codon, were prepared. The oligonucleotides were labelled at the 5' end using polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP (NEN). The specific activity of the oligonucleotides varied between 1,000 and 3,000 Ci mmol^{-1} oligonucleotide. A human genomic DNA library in bacteriophage λ[43] was screened using a modification of the *in situ* amplification procedure described originally by Woo *et al.*[44] and using tetramethylammonium chloride as the hybridization salt (see also refs 45–47; K.J. *et al.*, in preparation). Two independent phage (designated λHEPO1 and λHEPO2) hybridized to all three probes. DNA from λHEPO1 was digested to completion with *Sau*3A and subcloned into M13 for DNA sequence analysis using the dideoxy chain termination method[47]. Analysis of this DNA sequence revealed an open reading frame which precisely codes for the tryptic fragment used to deduce pool I. This open reading frame was contained in an 87-bp exon, bounded by potential splice acceptor and donor sites. Confirmation that λHEPO1 and λHEPO2 contain portions of the erythropoietin was obtained by identification, through further DNA sequencing of additional exons encoding amino-acid sequences corresponding to previously determined sequences of tryptic fragments of purified erythropoietin (see Figs 2, 3).



cccggagcc ggaccggggc caccgcgccc gctctgctccg acaccgcgcc
cccttgacag ccgcccctctc ctccaggccc gtgggggctgg ccctgcaccg ccgagcttcc cgggatgaggg ccccccggtgt
-27
MET GLY VAL HIS GLU CYS PRO
ggtcaccccgg cgcgccccag gtcgctgagg gaccccggcc aggcgcgggag ATG GGG GTG CAC GAA TGT CCT

ALA TRP LEU TRP LEU LEU LEU SER LEU LEU SER LEU PRO LEU GLY LEU PRO VAL LEU GLY
GCC TGG CTG TGG CTT CTC CTG TCC CTG CTG TCG CTC CCT CTG GGC CTC CCA GTC CTG GGC

1 SH 10 20
Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys
GCC CCA CCA CGC CTC ATC TGT GAC AGC CGA GTC CTG GAG AGG TAC CTC TTG GAG GCC AAG

SH 30 SH 40
Glu Ala Glu * Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu * Asn Ile Thr
GAG GCC GAG AAT ATC ACG ACG GGC TGT GCT GAA CAC TGC AGC TTG AAT GAG AAT ATC ACT

50 60
Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala
GTC CCA GAC ACC AAA GTT AAT TTC TAT GCC TGG AAG AGG ATG GAG GTC GGC CAG CAG GCC

70 80
Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
GTA GAA GTC TGG CAG GGC CTG GCC CTG CTC TCG GAA GCT GTC CTC CGG GGC CAG GCC CTC

* 90 100
Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser
TTG GTC AAC TCT TCC CAG CCG TGG GAG CCC CTG CAG CTG CAT GTG GAT AAA GCC GTC AGT

110 120
Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser
GGC CTT CGC AGC CTC ACC ACT CTG CTT CGC GCT CTG GGA GCC CAG AAG GAA GCC ATC TCC

130 140
Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys
CCT CCA GAT GCG GCC TCA GCT GCT CCA CTC CGA ACA ATC ACT GCT GAC ACT TTC CGC AAA

150 160
Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
CTC TTC CGA GTC TAC TCC AAT TTC CTC CGG GGA AAG CTG AAG CTG TAC ACA GGC GAG GCC

SH 166
Cys Arg Thr Gly Asp Arg
TGC AGG ACA GGG GAC AGA TGA ccaggtg tgtccacctg ggcatatcca ccacctcccc caccaacatt

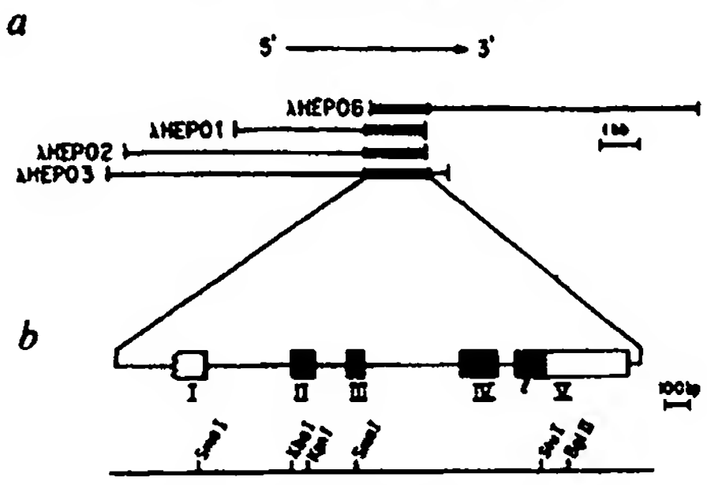
gcttgtgcca cacccctccc cgccactcct gaacccccgtc gaggggctct cagctcagcg ccagcctgtc
ccatggacac tccagtgcca gcaatgacat ctcaggggcc agaggaactg tccagagagc aactctgaga
tctaaggatg tcacagggcc aacttgaggg cccagagcag gaagcattca gagagcagct ttaaactcag
ggacagagcc atgctgggaa gacgcctgag ctcactcggc accctgcaaa atttgatgcc aggcacgct
ttggaggcga tttaccctgtt ttcgcaccta ccatcagggа cggatgacc tggcgaactt aggtggcaag
ctgtgacttc tccaggtctc acgggcatgg gcactcccct ggtggcaaga gcccccttga cacccgggtg
gtcggaacca tgaagacagg atgggggctg gcctctggct ctcatgggt ccaagttttg tgtattcttc
aacctcattg acaagaactg aaaccaccaa aaaaaaaaaaaa

Approximately 10 μg of human erythropoietin was purified from the urine of patients with aplastic anaemia and digested to completion with trypsin. The tryptic fragments were then purified by reverse-phase HPLC and subjected to microsequence analysis (ref. 16; see Fig. 1). We prepared highly degenerate synthetic oligonucleotides based on the amino-acid sequences and used these oligonucleotides to isolate the erythropoietin gene from a bacteriophage λ library of human genomic DNA (see Fig. 1).

The erythropoietin genomic clones were then used to determine whether human fetal liver is a potential source of messenger RNA for complementary DNA cloning, because erythropoietin is released from mouse[17], sheep[18] and human[19] fetal liver. Human fetal (20-week-old) and adult liver mRNAs were analysed by Northern blotting using as a probe a 95-nucleotide single-stranded fragment containing the 87-base pair (bp) exon described in Fig. 1. A strong signal was detected in fetal liver mRNA corresponding to an mRNA ~1,600 nucleotides in length (Fig. 1). An mRNA of identical size was detected weakly in adult liver mRNA and transcripts of ~2,000 nucleotides were detected weakly in both fetal and adult mRNA. The same probe was then used to isolate cDNA clones from a bacteriophage λ cDNA library constructed from the fetal liver mRNA[20].

The complete nucleotide and deduced amino-acid sequence for the largest of these clones (designated λHEPOFL13) is shown in Fig. 2. The erythropoietin coding information is contained in 579 nucleotides in the 5' half of the cDNA and encodes a hydrophobic 27-amino-acid leader peptide followed by the 166-amino-acid mature protein. The identification of the N-terminus of the mature protein is based on the N-terminal sequence of the protein secreted in the urine of patients with aplastic anaemia as determined originally by Goldwasser [21,22] and later confirmed (Fig. 1 and ref. 23). The amino acids underlined in Fig. 2 indicate the protein sequences obtained (see Fig. 1 legend) either from the N-terminus of intact erythropoietin or from purified tryptic fragments. The deduced amino-acid sequence agrees precisely with the protein sequence data, confirming that the isolated cDNA encodes human erythropoietin.

To demonstrate that biologically active erythropoietin could be expressed from the cloned cDNA, we performed transient expression experiments in COS cells[24]. The vector (p91023B) contains the adenovirus major late promotor, a simian virus 40 (SV40) polyadenylation sequence, an SV40 enhancer and origin of replication and the adenovirus virus-associated (VA) gene[25,26]. Erythropoietin cDNA was inserted into the p91023B vector downstream of the adenovirus major late promotor (Fig.



c

I

II

III

IV

V

Fig. 3 Structure of the erythropoietin gene. The relative sizes and positions of four independent genomic clones (λHEPO1, 2, 3, and 6) described in the text are illustrated by the overlapping lines in *a*. The thickened line indicates the position of the erythropoietin gene. The region containing the gene was sequenced completely from both strands using an exonuclease III generated series of deletions (C.S., unpublished observations) through this region. *b*, A schematic representation of five exons coding for erythropoietin mRNA. The precise 5' boundary of exon I is unknown (indicated by the broken box). The 5' boundary of exon I shown here is derived from λHEPOFL8, which has a 5'untranslated region 39 nucleotides longer than that of λHEPOFL13. The protein coding portion of the exons are darkened. *c*, Complete nucleotide sequences of the region. Exon sequences are given in capital letters; intron sequence in lower case letters. The location of exons I–V are indicated by the bars with numerals on the left. Because of difficulties in interpreting sequencing gel data from the very G+C-rich regions of exon I, the level of certainty for exon I sequence is reduced slightly.

2). After transfection of this construct into COS-1 cells, erythropoietin activity was detected by assays of the culture supernatant (Table 1).

Thus, the protein originally purified by Miyake *et al.*[27] and containing the N-terminus Ala-Pro-Pro-Arg. . . is erythropoietin (refs 21, 23; Fig. 2). Western blotting (using a polyclonal anti-erythropoietin antibody) indicates that erythropoietin produced in COS cells has a mobility on SDS-polyacrylamide gels identical to that of the native hormone prepared from human urine (data not shown).

As well as the clones described above (λHEPO1 and λHEPO2), two other genomic clones (λHEPO3 and λHEPO6) were isolated in subsequent screens of the human genomic library (Fig. 3*a*). Hybridization analysis of the cloned DNAs with oligonucleotide probes and with probes prepared from the erythropoietin cDNA clones positioned the erythropoietin gene in the 3.3-kilobase (kb) region in Fig. 3*a*. Complete sequence analysis of this region and comparison with the cDNA clones gave the map of intron and exon structure of the erythropoietin gene (Fig. 3*b*, *c*); the erythropoietin mRNA is encoded by at

Table 1 Assay for detection of erythropoietin activity

Assay method	Activity
In vitro CFU-E	2.0 ± 0.5 U ml^{-1}
In vitro ^{3}H-thymidine	3.1 ± 1.8 U ml^{-1}
In vivo exhypoxic mouse	1 U ml^{-1}
In vivo, starved rat	2.4 U ml^{-1}

The cDNA insert from λHEOPOFL13 was inserted into the vector p91023B (ref. 25) described in the text. Purified DNA (8 μg) was then used to transfect 5×10^6 M6 COS cells[37] using the DEAE-dextran method[25]; 12 h after transfection the cells were washed and exposed to media containing 10% fetal calf serum for 24 h. Cells were then changed to 4 ml serum-free media and collected 48 h later. *In vitro* biologically active erythropoietin was measured using either a colony-forming assay with mouse fetal liver cells as a source of erythroid colony-forming units (CFU-E)[38] or a ^{3}H-thymidine uptake assay using spleen cells from phenylhydrazine-injected mice[12]. Activities are expressed in units ml^{-1}, using a commercial, quantified erythropoietin (Toyobo, Inc.) as a standard. The sensitivities of the assays are ~25 mU ml^{-1}. *In vivo* biologically active erythropoietin was measured using either the hypoxic mouse[39] or the starved rat[40] method. The sensitivities of these assays are ~100 mU ml^{-1}. No activity was detected in either assay from mock-conditioned media. In subsequent experiments with the same vector, expression levels as high as 25 ± 3 U ml^{-1} (^{3}H-thymidine assay method) have been observed.

least five exons. Exons II, III, IV and parts of I and V contain the protein coding information, whereas the rest of exons I and V encode the 5′- and 3′-untranslated sequences, respectively. Exon I is 80% G+C and is surrounded by sequences equally G+C-rich. The CpG dinucleotide frequency in this region (~10%) is not significantly under-represented as it is in the remainder of the gene (~2%) and thus suggests a region of high methylation. The location of the actual cap site and the promoter region are not yet known.

The 166-amino-acid sequence deduced from the cDNA clones agrees precisely with our 102 amino acids of partial sequence of human urinary erythropoietin, including 25 residues at the N-terminus and 77 residues in 9 internal tryptic fragments. The sequence differs at four positions from the N-terminal sequences previously published[21-23], probably because of errors in interpretation or assignments in the original sequencing. The extent of identity between native human erythropoietin and the gene isolated here and the fact that we can detect only a single gene by genomic blotting with erythropoietin cDNA probes (data not shown) implies that the gene we have isolated is not a pseudogene or a closely related variant of the erythropoietin gene. If a second gene exists, it must be highly homologous over many kilobases to the gene described here.

We have assigned the N-terminus of the mature protein based on the N-terminus of the protein released into urine of individuals with aplastic anaemia, consistent with the hypothesis that the preceding 27 highly hydrophobic amino acids constitute a secretory leader peptide. One or more of the amino acids preceding the presumed mature terminus may be normally secreted with the remaining protein as a pro-form of erythropoietin, later processed to the native N-terminus. Amino-acid sequence analysis of tryptic fragments of urinary erythropoietin has not yet identified the fragment containing the C-terminal four amino acids (Thr-Gly-Asp-Arg; see Fig. 2). Thus, processing of erythropoietin may occur at the C-terminus and some or all of the final four amino acids encoded in the cDNA may be removed in this way. C-terminal sequencing of native erythropoietin or identification of the fragment will be necessary to answer this question.

There are four cysteines in the 166 amino acids of mature erythropoietin. Based on the sensitivity of the biological activity of erythropoietin to reducing agents (ref. 28 and T. Shimizu, personal communication), at least two of these residues must be involved in a disulphide bond.

In the mature protein there are three predicted sites of *N*-linked glycosylation (residues 24, 38 and 83) based on the consensus glycosylation site Asn-X-Ser/Thr[29]. Amino-acid sequence analysis suggests that the asparagines at residues 24 and 83 are glycosylated (data not shown) (residue 38 has not been examined). Native erythropoietin is highly glycosylated, displaying a complex, probably poly-antennary sugar structure[30]. The relative molecular mass (M_r) of the protein backbone deduced from the primary sequence is 18,398. As the reported M_rs for native erythropoietin determined by SDS gel electrophoresis are in the range 34,000–39,000 (refs 27, 31), nearly one-half of the apparent M_r of erythropoietin must be contributed by the sugar side chains. Whether any of the glycosylation is the result of *O*-linked glycosylation is unknown. The terminal sialic acid residue(s) of native erythropoietin is required for full *in vivo* biological activity but is not necessary for *in vitro* activity[32]. This effect may result from enhanced clearance of asialylated erythropoietin from the circulation by the liver[33]. The biological activity of a completely unglycosylated erythropoietin may now be assessable using a recombinant system.

Lee-Huang[34] recently reported the isolation of an erythropoietin cDNA clone from mRNA of a human kidney carcinoma. As no sequence information was provided, we are unable to compare the erythropoietin clones described here with the cDNA clone of Lee-Huang[34]. Fyhrquist *et al.*[35] have suggested that renin substrate (angiotensinogen) may be the erythropoietin precursor. Our results argue against a large precursor and comparison of the human erythropoietin amino-acid sequence with the rat angiotensinogen protein sequence[36] reveals no regions of homology and further argues against any relationship between the two polypeptides. Finally, extensive comparison of the erythropoietin amino-acid and cDNA sequence with sequences contained in both the National Biomedical Research Foundation and Genbank data bases has revealed no significant homology with any published sequence.

We thank Dr Judith Sherwood for the anti-erythropoietin antibody, Dr John Tooze for the fetal liver cDNA library, Drs Peter Dukes and Masayoshi Ono for the *in vivo* biological assays, Dr Eugene L. Brown for helpful discussions on the selection of oligonucleotide probes, John Brown, Tatjana Loh, Chris Bassler, Pat Murtha, Louise Wasley, Richard Wright, Evan Beckman, Ann Leary, Tom Gesner, Jane Aghajanian and Lisa Mitsock for technical support, Elizabeth Orr for help with the computer analysis, Joyce Lauer for improvements to the manuscript, Marybeth Erker for typing the manuscript and Dr Robert Kamen and Gabriel Schmergel for their support and encouragement. This project was supported by Chugai Pharmaceuticals, Japan.

Received 17 December 1984; accepted 30 January 1985.

1. Sherwood, J. B. & Goldwasser, E. *Endocrinology* 103, 866-870 (1978).
2. Hammond, D. & Winnick, S. *Ann. N.Y. Acad. Sci.* 230, 219-227 (1974).
3. Jacobson, L. O., Goldwasser, E. Fried, W. & Plzak, L. F. *Trans. Ass. Am. Phys.* 10, 305-317 (1957).
4. Krantz, S. B. & Jacobson, L. O. thesis, Univ. Chicago (1970).
5. Fried, W. *Blood* 40, 671-677 (1972).
6. Naughton, B. A. *et al.*, *Science* 196, 301-302 (1977).
7. Lucarelli, G. P., Howard, D. & Stohlman, F. Jr *J. clin. Invest.* 43, 2195-2203 (1964).
8. Zanjani, E. D., Poster, J., Burlington, H., Mann, L. I. & Wasserman, L. R. *J. Lab. clin. Med.* 89, 640-644 (1977).
9. Fisher, J. *Proc. Soc. exp. Biol. Med.* 173, 289-305 (1983).
10. Krantz, S. B., Gallfen-Lartigue, O. & Goldwasser, E. *J. biol. Chem.* 238, 4085-4090 (1963).
11. Dunn, C. D., Jarvis, J. H. & Greenman, J. M. *Expl Hemat.* 3, 65-78 (1975).
12. Krystal, G. *Expl Hemat.* 11, 649-660 (1983).
13. Krane, N. *Henry Ford Hosp. Med. J.* 31, 177-181 (1983).
14. Anagnostou, A., Barone, J., Vedo, A. & Fried, W. *Br. J. Hemat.* 37, 85-91 (1977).
15. Eschbach, J., Mladenovic, J., Garcia, J., Wahl, P. & Adamson, J. *J. clin. Invest.* 74, 434-441 (1984).
16. Hewick, R. M., Hunkapiller, M. E., Hood, L. E. & Dreyer, W. J. *J. biol. Chem.* 256, 7990-7997 (1981).
17. Zanjani, E. D., Ascensao, J. L., McGlave, P. B., Banisadre, M. & Ash, R. C. *J. clin. Invest.* 67, 1183-1188 (1981).
18. Gruber, D. F., Zucali, J. R. & Mirand, E. A. *Expl Hemat.* 5, 392-398 (1977).
19. Congote, L. F. *J. Steroid Biochem.* 8, 423-428 (1977).
20. Toole, J. J. *et al. Nature* 312, 342-347 (1984).
21. Goldwasser, E. *Blood* 58, Suppl. 1, 13 (abstr.) (1981).
22. Sue, J. M. and Sytkowdki, A. J. *Proc. natn. Acad. Sci. U.S.A.* 80, 3651-3655 (1983).
23. Yanagawa, S. *et al. J. biol. Chem.* 259, 2707-2710 (1984).
24. Gluzman, Y. *Cell* 23, 175-182 (1981).
25. Wong, G. C. *et al. Science* (in the press).
26. Kaufman, *Proc. natn. Acad. Sci. U.S.A.* (in the press).
27. Miyake, T., Kung, C. & Goldwasser, E. *J. biol. Chem.* 252, 5558-5564 (1977).
28. Sytkowski, A. *Biochem. biophys. Res. Commun.* 96, 143-149 (1980).
29. Wagh, P. V. & Bahl, O. P. *CRC crit. Rev. Biochem.* 307-377 (1981).
30. Murphy, M. & Miyake, T. *Acta. Haemat. Jap.* 46, 1380-1396 (1983).
31. Wang, F. F., Kung, C. K.-H. & Goldwasser, E. *Fedn Proc.* 42, 1872 (abstr.) (1983).

32. Lowy, P., Keighley, G. & Barnook, H. *Nature* **185**, 102-103 (1960).
33. VanLenten, L. & Ashwell, G. *J. biol. Chem.* **247**, 4633-4640 (1972).
34. Lee-Huang, S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2708-2712 (1984).
35. Fyrquist, F., Rosenlof, K., Gronhagen-Riska, C., Hortling, L. & Tikkanen, I. *Nature* **308**, 649-652 (1984).
36. Ohkubo, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 2196-2200 (1983).
37. Horowitz, M., Cepko, C. & Sharp, P. A. *J. molec. appl. Genet.* **2**, 147-149 (1983).
38. Bersch, N. & Golde, D. W. in *In Vitro Aspects of Erythropoiesis* (Murphy, M. J.) 252-253 (Springer, New York, 1978).
39. Cotes, P. M. & Bangham, D. R. *Nature* **191**, 1065-1068 (1961).
40. Goldwasser, E. & Gross, M. *Meth. Enzym.* **37**, 109-121 (1975).
41. Derman, E. *et al. Cell* **23**, 731-739 (1981).
42. Anderson, S. & Kingston, I. B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6836-6842 (1983).
43. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157-1174 (1978).
44. Woo, S. L. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3688-3691 (1978).
45. Melchior, W. B. & von Hippel, P. H. *Proc. natn. Acad. Soc. U.S.A.* **70**, 298-302 (1973).
46. Orosz, J. M. & Wetmur, J. G. *Biopolymers* **16**, 1183-1199 (1977).
47. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
48. Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).

Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter

Joan T. Odell, Ferenc Nagy & Nam-Hai Chua

Laboratory of Plant Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10021-6399, USA

Although promoter regions for many plant nuclear genes have been sequenced, identification of the active promoter sequence has been carried out only for the octopine synthase promoter[1]. That analysis was of callus tissue and made use of an enzyme assay. We have analysed the effects of 5′ deletions in a plant viral promoter in tobacco callus as well as in regenerated plants, including different plant tissues. We assayed the RNA transcription product which allows a more direct assessment of deletion effects. The cauliflower mosaic virus (CaMV) 35S promoter provides a model plant nuclear promoter system, as its double-strand DNA genome is transcribed by host nuclear RNA polymerase II from a CaMV minichromosome[2]. Sequences extending to −46 were sufficient for accurate transcription initiation whereas the region between −46 and −105 increased greatly the level of transcription. The 35S promoter showed no tissue-specificity of expression.

The 35S promoter region was isolated as a *Bgl*II fragment extending from −941 to +208 with respect to the transcription start site mapped for the 35S RNA found in CaMV-infected turnip leaves[3]. The polyadenylation site for the 19S and 35S CaMV transcripts located at +180 (ref. 3) was deleted, as described in Fig. 1 legend, to eliminate any possible processing signals in the promoter fragment. A 3′ deleted promoter fragment extending to +9 was deleted at its 5′ end (see Fig. 1) and fragments extending to −343, −168, −105 and −46 were chosen for analysis.

An abbreviated human growth hormone gene (*hgh*)[4] was added as a test gene downstream to the 35S promoter deletion fragments. Information on plant cell recognition of animal gene splice and 3′ polyadenylation signals obtained from analysis of *hgh* RNA transcribed in transformed plant cells will be presented elsewhere (A. Hunt, N. Chu, J.T.O., F.N. and N.-H.C., in preparation). The 35S promoter–*hgh* chimaeric gene was inserted in the pMON178 tumour-inducing (Ti)-plasmid vector, a derivative of pMON120 (ref. 5). Included in this vector is the nopaline synthase (NOS) promoter placed 5′ to the neomycin phosphotransferase-II (*npt*-II) coding region (NOS promoter–*npt*-II gene), which is co-transferred with the 35S promoter–*hgh* gene into the tobacco genome and provides an internal standard for comparison of the activities from different 35S promoter deletion fragments.

Following tri-parental matings[5,6], *Agrobacterium tumefaciens* containing both chimaeric genes was used to infect SR1 *Nicotiana tabacum* cells by wounding[5] and co-cultivation[5,7].

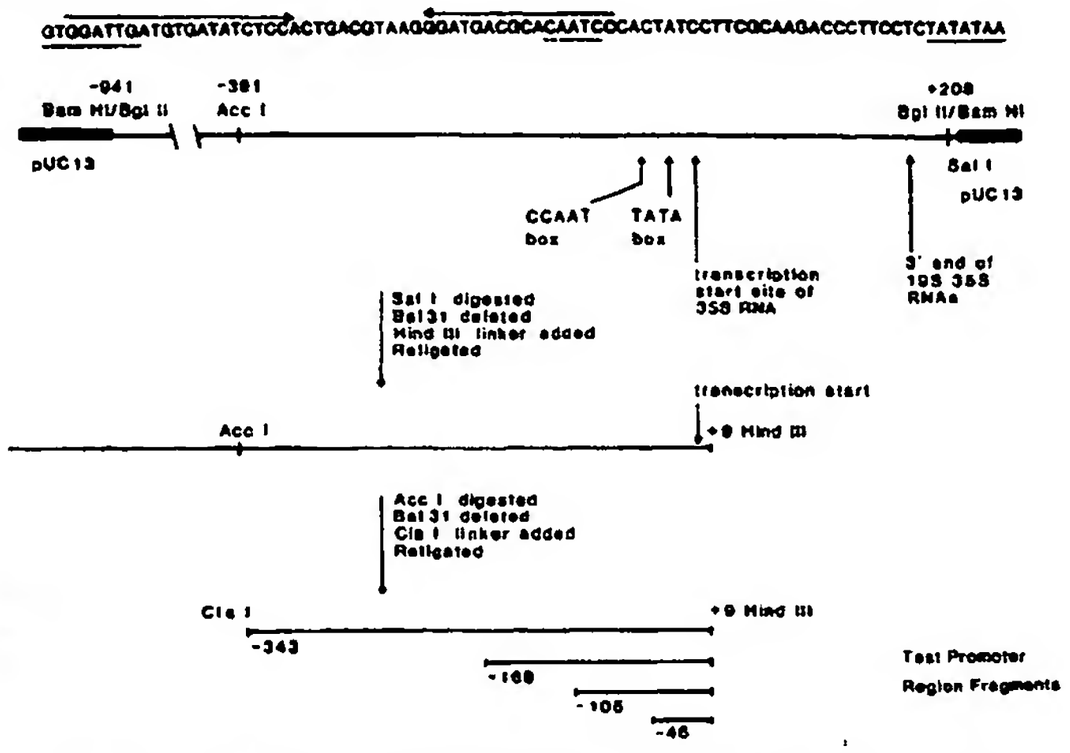


Fig. 1 Construction of 35S promoter region fragments. A 1.15-kb *Bgl*II fragment was subcloned from pCS101, a clone containing the entire Cabb-S CaMV genome[3], into the *Bam*HI site of pUC13. The resulting plasmid was linearized at the *Sal*I site in the pUC13 polylinker next to the 3′ end of the promoter fragment, digested with *Bal*31 exonuclease[11], ligated to *Hin*dIII linkers and recircularized. Clones were analysed for the extent of 3′ deletion by polyacrylamide gel sizing of the *Acc*I/*Hin*dIII fragments and finally by dideoxy sequencing[12] of subclones in pUC using the universal primer. The plasmid containing a 3′ deletion fragment with the *Hin*dIII linker at +9 was linearized with *Acc*I (site at −391), digested with *Bal*31 exonuclease, ligated to *Cla*I linkers and recircularized. Clones were analysed for the extent of 5′ deletion by polyacrylamide gel sizing of the *Cla*I/*Hin*dIII fragment, followed by dideoxy sequencing of subclones in pUC using either the universal primer or primer generation by exonuclease III digestion[13]. Above is the sequence of the −105 to −25 region of the 35S promoter[14] with TATA-box, CAAT-box, inverted repeat and core enhancer sequence regions marked.

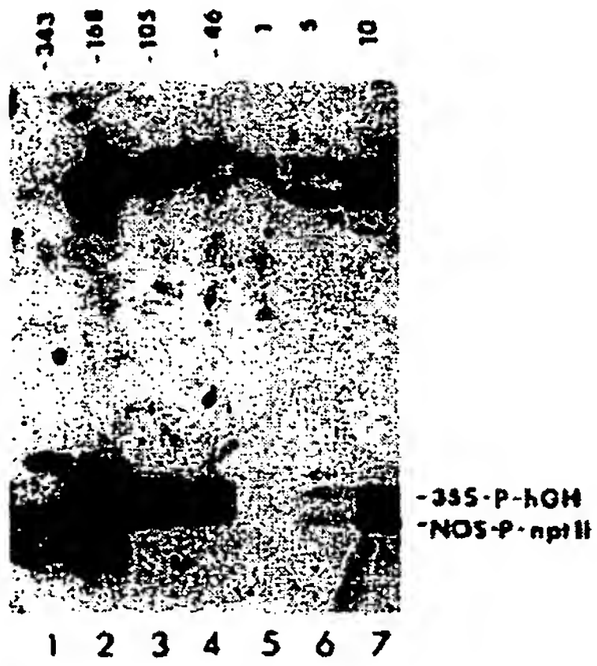


Fig. 2 Southern blot analysis of DNA from transformed tobacco calli. DNA was prepared, digested with *Eco*RI, electrophoresed on a 0.7% agarose gel and blotted onto a nitrocellulose filter[15]. A plasmid constructed to serve as the hybridization probe contains a *Bam*HI/*Sma*I *hgh* gene fragment and a *Bam*HI/*Bgl*II *npt*-II gene fragment cloned into pUC12 (GH-Neo24). The plasmid was nick translated[16] and hybridized to the Southern blot by the method of Thomashow *et al.*[17]. The following samples contain 15 μg of calli DNA transformed with: lane 1, −343 35S promoter–*hgh*; lane 2, −168 35S promoter–*hgh*; lane 3, −105, 35S promoter–*hgh*; lane 4, −46 35S promoter–*hgh*. Reconstructions of the NOS promoter–*npt*-II gene and 35S promoter–*hgh* gene copy numbers contain 15 μg of control untransformed plant DNA mixed with different amounts of the pMON178 plasmid containing the −105 35S promoter–*hgh* gene: lane 5, 17 pg = 1 copy; lane 6, 85 pg = 5 copies; lane 7, 170 pg = 10 copies. The bands near the top of the filter in lanes 1–4 result from hybridization of the pBR322 sequences in the GH-Neo24 probe plasmid to pBR322 sequences in the integrated pMON178 Ti vector. In lanes 5–7 the upper bands are derived from other regions of the pMON178 plasmid.